*International Research Journal of Innovations in Engineering and Technology (IRJIET)*

**ISSN (online): 2581-3048**
**Volume 8, Issue 4, pp 330-335, April-2024**
https://doi.org/10.47001/IRJIET/2024.804052

# Android Application for Sign Language to Text and Speech Conversation

**[1]Prof. Sagar S. Mane, [2]Swapnil Ohol, [3]Adarsh Nayse, [4]Akshay Patil, [5]Rohit Teke**

[1]Professor, Department of Computer Engineering, NBN Sinhgad Technical Institutes Campus, Pune, Maharashtra, India
[2,3,4,5]Student, Department of Computer Engineering, NBN Sinhgad Technical Institutes Campus, Pune, Maharashtra, India

***Abstract* - Living in the 21st century we are still facing the problem of communicating deaf and dumb disability people. As we know they having a particular set of sign languages they make gestures through hands to convey their message to the opposite person but in the major of cases sign language becomes a barrier for a normal person because only the trained person can understand it by decoding the sign gesture. So this project deals with an application which will translate some common sign language gesture through image processing by detecting the hand gestures via mobile camera and convert into a text message format and send it through the social media site so the person with no knowledge of sign language gestures can also understand it and communicate with the opposite person. The proposed system will fully be based on the software and try to reduce the hardware burdens of the existing one. It will help to reduce the language barrier from the deaf and dumb community.**

***Keywords:*** Android Application, Sign Language, Image Processing, Mobile camera.

## I. INTRODUCTION

Android application have shown a dramatic improvement in their functionality to a point where it is now possible to have cellular phone execute Java programs. As a result, cellular users throughout the world are now able to read and write email, browse web pages and play java games using their cellular phones. This trend has promoted as to propose the use of android application for better communication. Before SMS/MMS, deaf people rarely used mobile phones. Now texting allows deaf people remotely to communicate with both deaf and hearing parties. Mobile video chat may one day replace texting, but only for conversations between hearing callers, not for those between deaf and hearing callers. It is an application in which an image movement will repeat everything we say in a high-pitched voice. Without dialing number, we can use this application.

This project deals an alternative for gesture detection using image processing technique between deaf people which overcomes the above technique and paves the way for the communication between deaf and normal people in their daily activities using sign language and video relay service. Video technology continues to improve and one day may be the preferred means of mobile communication among the deaf. It allows deaf, hard-of-hearing and speech impaired individuals to communicate over video or other technology with hearing people in real-time, via a sign language interpreter. The idea behind SE (Signed English) and other signing system parallel to English is the deaf people will learn English better if they are exposed.

The sign language provides video by improving small-screen mobile communication among the deaf. There are mainly three parts:

- Speech-Recognition Engine
- Database and
- Recognized Text

## II. LITERATURE SURVEY

The purpose of the Literature Survey is to give the brief overview and also to establish complete information about the reference papers. The goal of Literature Survey is to completely specify the technical details related to the main project in a concise and unambiguous manner.

In different approaches have been used by different researchers for recognition of various hand gestures which were implemented in different fields [1]. The whole approaches could be divided into three broad categories:

- Hand segmentation approaches
- Feature extraction approaches and
- Gesture recognition approaches.

All the available systems are not portable and not affordable to poor people. This paper introduces new android application which will detect the Indian sign language via mobile camera and converts into corresponding text or voice output. Now our system provides 65% of correct predicting and we are working on improving its efficiency. Hence we took the idea of implementing the gesture video with the help of hand speak technology which helps the deaf people to view their relevant sign language video based on the text given as

 

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 330-335, April-2024

https://doi.org/10.47001/IRJIET/2024.804052

input. We include the idea of providing the link to the application which helps in extracting the video. It proves its maximum efficiency.

Sign language is used as a communication medium among deaf and dumb people to convey the message with each other [2]. In order to bridge the gap in communication among deaf, dumb community and normal community, lot of research work has been carried out to automate the process of sign language interpretation with the help of image processing and pattern recognition techniques. This paper proposes optimized approaches of implementing the famous Viola Jones algorithm with LBP (Local Binary Pattern) features for hand gesture recognition which will recognize Indian sign language gestures in a real time environment. An optimized algorithm has been implemented in the form of an android application and tested with real time data. This implemented algorithm is not a robust and real time. Hence we are using the already recorded video stored in a cloud storage which is considered to be the easiest way of interpreting the users input in relevant manner. This above algorithm does not prove its efficiency in any sort of background but our project overcomes this issue to the larger extent.

A number of developing countries continue to provide educational services to students with disabilities in "segregated" schools. Also all students, regardless of their personal circumstances, have a right of access to and participation in the education system, according to their potential and ability [3]. However, with the rapidly growing population and increasing number of people with blindness along with other disabilities, need for use of technology in the field of education has become imminent. In this project, through the use of speech technology, attempts to provide solutions for some of these issues by creating an interactive system. We took the idea of using voice over text technology from the above proposed system because on considering the deaf people, they either have speech ability or be a dumb which again depends on their birth. It will be a revolutionary change that willbenefit hearing impaired people, boost their confidence and put them with regular people.

For the past several decades, designers have processed speech for a wide variety of applications ranging from mobile communications to automatic reading machines [4]. Speech has not been used much in the field of electronics and computers due to the complexity and variety of speech signals and sounds. Our speech-to-text system directly acquires and converts speech to text. It can supplement other larger systems, giving users a different choice for data entry. A speech-to-text system can also improve system accessibility by providing data entry options for blind, deaf, or physically handicapped users. Voice SMS is an application developed in this work that allows a user to record and convert spoken messages into SMS text message. User can send messages to the entered phone number. Speech recognition for Voice uses a technique based on hidden Markov models (HMM - Hidden Markov Model). It is currently the most successful and most flexible approach to speech recognition. Using the speech recognizer, which works over the Internet, allows much faster data processing.

Text-to-speech (TTS) convention [5] transforms linguistic information stored as data or text into speech. It is widely used in audio reading devices for blind people now a day. In the last few years however, the use of text-to-speech conversion technology has grown far beyond the disabled community to become a major adjunct to the rapidly growing use of digital voice storage for voice mail and voice response systems. This paper presents a method to design a Text to Speech conversion module by the use of Matlab by simple matrix operations.

## III. PROPOSED SYSTEM

**Proposed System:**

The proposed system architecture has been shown in Fig. 1.1. This system will be built for the normal people those which will be communicating with deaf & dumb community. Users need to hold their android phone's camera pointing towards the disabled person and start the video. The trained model at the backend will process the frames of video and recognize the hand gesture made by the disabled person. Further the semantic of signs will be converted to Speech/Voice.

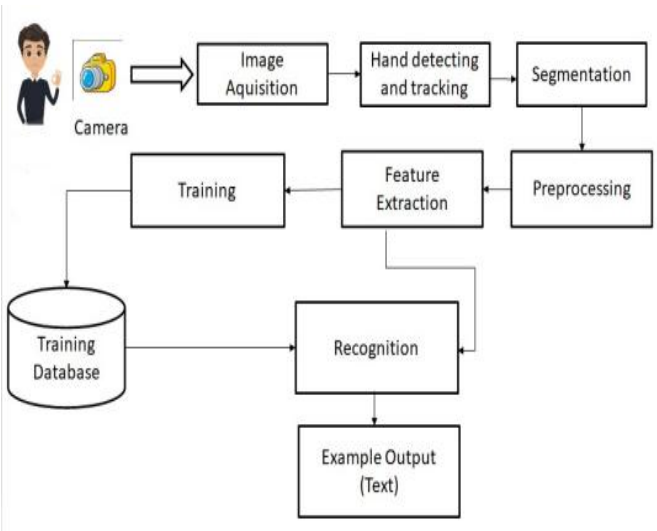


**Figure 1: Proposed System**

The purpose of this project is to develop an android mobile application that uses gesture recognition to understand

 

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 330-335, April-2024

https://doi.org/10.47001/IRJIET/2024.804052

American Sign Language. The android application should use image processing techniques to distinguish the hand from the background and then identify the fingertips to determine the gestures. The recognition is limited to only static features and hence is only used for detecting static American Sign Language alphabets. The android application should understand the gestures and show it in text form.

The project has the following main tasks:

- Research about different Image Processing techniques
- Extract skin colored pixels and convert every frame to binary image
- Extract hand pixels from rest of the hand be extracting the largest contour
- Identify fingertips and centre of the hand from the largest contour
- Using the distances between the fingers and centroid, recognize the gesture
- Display the gesture in textual form to show which alphabet is shown as a gesture
- Convert text into audio form

## IV. RESULTS AND DISCUSSION

### CNN Algorithm

### A) Neural Network

The neural network consists of a number of algorithms that allow them to recognize fundamental links in an entire series of data, and this process imitates the way our brains work. A neuron network is defined as any system of neurons, either natural or artificial. NRPs are capable of adapting to changes in input, so that they can achieve the best possible result and do not require a revision of their output criteria. In the field of recognition systems, the concept of neural networks that has its origins in artificial intelligence is rapidly gaining popularity. The neural network is similar to the human brain's neural network. A “neuron” in a neural network is a mathematical function that gathers and classifies information according to a specific architecture. The network bears a strong resemblance to statistical methods such as curve fitting and regression analysis. A neural network contains layers of interconnected nodes. Each node is apperception and is similar to multiple linear regressions. The perception feeds the signal produced by a multiple linear regression into an activation function that may be nonlinear.

### B) Convolutional Neural Network

Another approach is taken by the Convolution Neural Network that replicates how we view our environment from a visual perspective. When we see an image, it's automatically divided into a bunch of smaller sub-images that are analyzed one at a time. We will process and analyze the image after we have assembled these sub-images. Convolution layers are capable of implementing this principle. To do this, we'll define a filter that will determine the size of the partial images to be looked at and step length which determines what number of pixels are left in between calculations. We have significantly decreased the dimensionality of that image by taking this step. The convolution allayers the foundation for a convolutional network. While the convolutional layer may be followed by other convolutional layers or pooling layers, the fully connected layer is the final layer. The CNN gets more complex as the layers increase in complexity, meaning that a greater proportion of the image is identified. Small features, e.g. colors and edges, tend to be the focus of older layers.

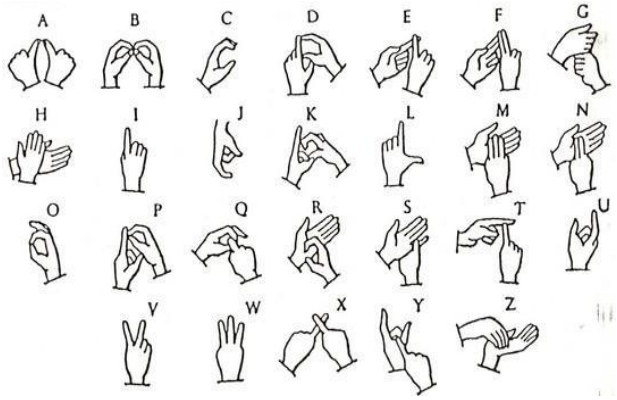


**Figure 2: CNN working**

### C) Methodology

***Image Acquisition:***

It is the process of taking an image out of a source usually one that is hardware-based in order to process it. The hardware-based component of our concept is Web Camera. It is the initial stage in the workflow sequence since a picture is necessary for all processing to be done. The resulting image has not undergone any kind of processing.

***Segmentation:***

Segmentation is the process of taking items or signs out of the background of an image that has been taken. The segmentation process makes use of edge detection, context subtraction, and skin-color detection. Recognizing gestures requires the detection and segmentation of hand motion and location.

***Features Extraction:***

Preprocessed images are used to extract predefined features, such as form, contour, geometrical feature (position, angle, distance, etc.), color feature, histogram, and others,

 

ISSN (online): 2581-3048
Volume 8, Issue 4, pp 330-335, April-2024
https://doi.org/10.47001/IRJIET/2024.804052

which are then utilized for sign identification or classification. A step in the dimensionality reduction process that separates and arranges a sizable amount of raw data is feature extraction. Lowered to more manageable, smaller classrooms Processing would be easier as a result. The most significant aspect is the sheer quantity of variables present in these enormous data sets. This data requires a significant amount of processing power to process. So, by choosing and combining variables into functions, function extraction helps to extract the optimal feature from enormous data sets. These features accurately and uniquely describe the actual data collection process, and they are also very user-friendly.

***Preprocessing:***

Preprocessing techniques such as erosion, dilation, and Gaussian smoothing are applied to each image frame in order to remove noise. Converting a color image to grayscale results in a smaller image. Grayscale picture conversion is a popular way to minimize the quantity of data that needs to be processed. Preprocessing goes through the following phases:

*a) Morphological Transform (Morphological Transform):*

To produce an output image with a comparable size, morphological processes employ a structural feature on the input image. To find the value of each pixel in the output image, it compares the matching pixel in the input image with its neighbors. Morphological alterations come in two flavors: erosion and dilation.

Dilation: The output pixel's value is the maximum value of all the nearby pixels. If every pixel in a binary image has the value 1, then that pixel is set to 1. Morphological dilatation closes tiny gaps and makes artifacts more visible.

Erosion: The value of the o/p pixel is the lowest of all the nearby pixels. In a binary image, a pixel is set to 0 if every one of its neighbors is also 0. Morphological erosion removes minor artifacts, leaving behind larger objects.

*b) Blurring:*

Blurring occurs when an image is passed through a low-pass filter. In computer vision, the term "low-pass filter" describes a technique for removing noise from an image while preserving the integrity of the remaining portion. Before moving on to more complex tasks, including edge detection, a blur is a basic process that needs to be finished.

*c) Thresholding:*

Thresholding is a type of image segmentation where an image's pixels are altered to facilitate picture interpretation. The act of turning a grayscale or color image into a binary—basically, black and white—is known as thresholding. Thresholding is most frequently used to choose regions of interest in an image while disregarding the parts we don't care about.

*d) Recognition:*

In this instance, classifiers will be used. The techniques or algorithms used to interpret the signals are called classifiers. Convolution Neural Network (CNN), and Principle Component Analysis (PCA) are a few popular classifiers that recognize or comprehend sign language. CNN will be used as the classifier in this research, nevertheless. CNNs are utilized for picture recognition and classification due to their high precision. Utilizing a hierarchical paradigm, the CNN creates a network akin to a funnel before producing a fully-connected layer that processes the output and connects all of the neurons.

*e) Text output:*

Recognizing and translating into text different body postures and movements, as well as comprehending human behavior.

**Result:**

It can be inferred from the results that the validation test of the Sign Language Recognition System was considered legitimate that the system is a valuable tool for translating sign into sign language.

We have developed a model that can recognize gestures in Indian sign language and provide textual and Numerical output. Three possibilities were created in the sign predictor. Such as life prediction, image loading, and image classification. If the user wishes to create a forecast based on photographs, we offer real-time gesture detection using life prediction. To facilitate this process, we have included the opportunity for the user to hand movement. We obtained 98% accuracy for both the alphabet and numbers and 96% accuracy for words when we trained the model. When we provide the system with input, we need new light for the system to function at its best. 90% of the data were utilized for training, and 10% were used for testing. More data might be used for testing and training, which would improve the accuracy and output of the system.

 

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 330-335, April-2024

https://doi.org/10.47001/IRJIET/2024.804052

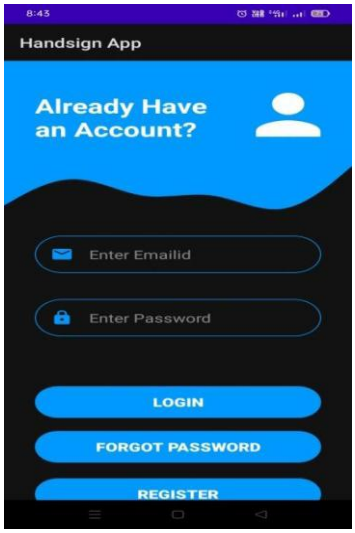


**Figure 3: Home page**

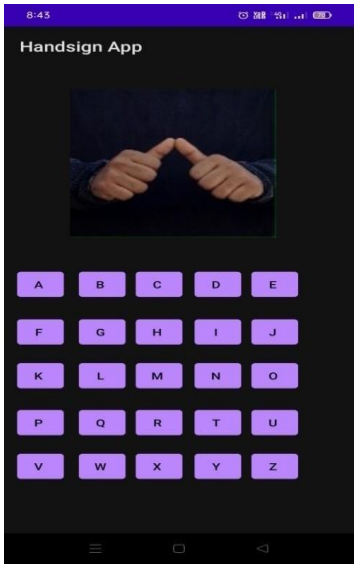


**Figure 4: Sign Recognition**

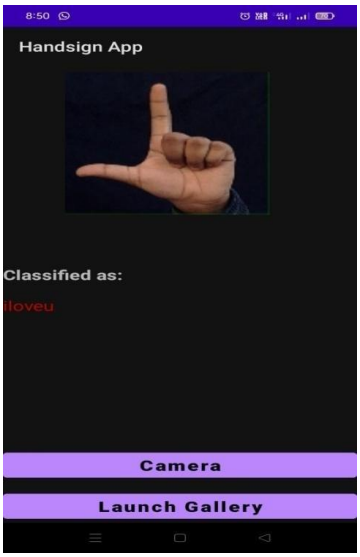


**Figure 5: Sign Result**

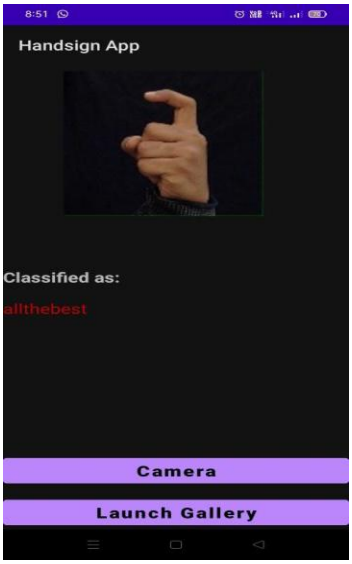


**Figure 6: Sign Identification**

## V. CONCLUSION

Sign Language Translator will act as strong bridge of communication between normal people and deaf & Community. This application will lead people to instantly interpret the sign language without any need third person. This application will help society to communicate the deaf & dumb community and generate a bond between them. We are focusing that this application will help to interact with deaf and dumb person at places such as Shops, Hospitals, Police headquarters etc. Further implementation of our system will enable text and audio support with gesture recognition.

## REFERENCES

[1] Raghavendhar Reddy.B, Mahender.E, "Speech to Text Conversion using Android Platform", International Journal of Engineering Research and Applications (IJERA) Vol. 3, Issue 1, pp.253-258, January -February 2013.

[2] Sangeetha, K. and Barathi Krishna, L. "Gesture Detection for Deaf and Dumb People " International Journal of Development Research Vol. 4, Issue, 3, pp. 749-752, March, 2014.

[3] Shanmukha Swamy, Chethan M P, Mahantesh Gatwadi, "Indian Sign Language Interpreter with Android Implementation", International Journal of Computer Applications (0975 – 8887) Volume 97– No.13, pp. 36-41, July 2014.

[4] Sinora Ghosalkar, Saurabh Pandey, Shailesh Padhra, Tanvi Apte, "Android Application on Examination Using Speech Technology for Blind People", International Journal of Research in Computer and Communication Technology, Vol 3, Issue3, March-2014.

 

ISSN (online): 2581-3048
Volume 8, Issue 4, pp 330-335, April-2024
https://doi.org/10.47001/IRJIET/2024.804052

[5] Tapas Kumar Patra, Biplab Patra, Puspanjali Mohapatra, "Text to Speech Conversion with Phonematic Concatenation", International Journal of Electronics Communication and Computer Technology (IJECCT) Volume 2 Issue 5 , pp.223-226, September 2012.

[6] Hand Gesture Recognition and Voice Conversion for Deaf and Dumb, Rupesh Prajapati1, Vedant Pandey2, Nupur Jamindar3, Neeraj Yadav4, Prof. Neelam Phadnis, Department of Computer Engineering, Shree L.R. Tiwari College of Engineering, Maharashtra, India 16).

[7] Design of Communication Interpreter for Deaf and Dumb Person, Pallavi Verma1, Shimi S. L. 2, Richa Priya darshani 3 Electrical and Electronics Department, Amity University, Greater Noida, Uttar Pradesh, India 2Assistant Professor, NITTTR, Chandigarh, India.

[8] Design and Implementation of A Sign to Speech/Text System for Deaf and Dumb People, DalalAbdulla, Shahrazad Abdulla and Rameesa Manaf Electrical and Computer Engineering Department University of Sharjah Sharjah 27272, UAE.

[9] Sign Language Recognition, Anup Kumar, Karun Thank achan and Mevin M. Dominic Department of Computer Science and Engineering National Institute of Technology, Calicut, India 673601.

[10] Image Caption Generation using Deep Learning Technique, Chetan Amritkar Department of EnTC Vishwakarma Institute of Technology Pune, India.

## AUTHORS BIOGRAPHY

**Rohit Teke,**
Student, Department of Computer Engineering, NBN Sinhgad Technical Institutes Campus, Pune, Maharashtra, India.

**Akshay Patil,**
Student, Department of Computer Engineering, NBN Sinhgad Technical Institutes Campus, Pune, Maharashtra, India.

**Adarsh Nayse,**
Student, Department of Computer Engineering, NBN Sinhgad Technical Institutes Campus, Pune, Maharashtra, India.

**Swapnil Ohol,**
Student, Department of Computer Engineering, NBN Sinhgad Technical Institutes Campus, Pune, Maharashtra, India.

**Citation of this Article:**

Prof. Sagar S. Mane, Swapnil Ohol, Adarsh Nayse, Akshay Patil, Rohit Teke, "Android Application for Sign Language to Text and Speech Conversation", Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 8, Issue 4, pp 330-335, April 2024. Article DOI https://doi.org/10.47001/IRJIET/2024.804052

*******

 